

CLAIMS

1. A network system, comprising:
 - a first network node, comprising:
 - an input for receiving a packet;
 - a buffer, coupled to the input and for storing the packet;
 - 5 circuitry for detecting when a number of packets stored in the buffer exceeds a buffer storage threshold; and
 - circuitry, responsive to a detection by the circuitry for detecting that the number of packets stored in the buffer exceeds the buffer storage threshold, for issuing a pause message along an output to at least a second network node;
 - 10 wherein the pause message indicates a message ingress address and a message egress address, the message ingress address and the message egress address corresponding to a network ingress address and a network egress address in a congestion-causing packet received by the first network node; and
 - wherein the pause message commands the second network node to discontinue,
 - 15 for a period of time, transmitting to the first network node any packets that have the message ingress address and the message egress address.

2. The system of claim 1:

wherein the first network node further comprises:

circuitry for designating a plurality of virtual space regions in the buffer;

and

5 circuitry for detecting packet congestion in at least one congested virtual space region in the plurality of virtual space regions;

wherein each virtual space region in the plurality of virtual space regions corresponds to a single network ingress address and a single network egress address;

10 wherein the buffer storage threshold represents a global threshold indicating a packet occupancy for all packets in all of the plurality of virtual space regions; and

wherein the circuitry for issuing the pause message is further responsive to issuing the pause message in response to detection of the at least one congested virtual space region.

3. The system of claim 2 wherein the circuitry for detecting packet congestion detects in response to packet occupancy in the at least one congested virtual space region exceeding a region threshold corresponding to the at least one congested virtual space region.

4. The system of claim 3 wherein the pause message indicates the single network ingress address and a single network egress address corresponding to the at least one congested virtual space region.

5. The system of claim 3 wherein the pause message indicates the period of time.

6. The system of claim 5 wherein the first network node further comprises circuitry for determining the period of time in response to an amount of a number of packets in the buffer and exceeding the region threshold of the at least one congested virtual space region.

7. The system of claim 6:
wherein the buffer comprises a first buffer;
wherein the first network node further comprises a second buffer;
wherein the first buffer is for storing a first class of packets; and
5 wherein the second buffer is for storing a second class of packets different from the first class of packets.

8. The system of claim 7 wherein the second buffer is not divided into a plurality of virtual space regions.

9. The system of claim 7 wherein the circuitry for designating is further for designating a plurality of virtual space regions in the second buffer.

10. The system of claim 9:
wherein the buffer storage threshold comprises a first buffer storage threshold;
wherein the circuitry for detecting is further for detecting when a number of packets stored in the second buffer exceeds a second buffer storage threshold; and
5 wherein the circuitry for issuing a pause message along an output to at least a second network node is further responsive to detection by the circuitry for detecting that the number of packets stored in the second buffer exceeds the second buffer storage threshold.

11. The system of claim 10 and further comprising the second network node.

12. The system of claim 10 wherein the second network node is directly-connected to the first network node.

13. The system of claim 1:
wherein the buffer comprises a first buffer;
wherein the first network node further comprises a second buffer;
wherein the first buffer is for storing a first class of packets; and
5 wherein the second buffer is for storing a second class of packets different from the first class of packets.

14. The system of claim 13 wherein the second buffer is not divided into a plurality of virtual space regions.

15. The system of claim 13 wherein the circuitry for designating is further for designating a plurality of virtual space regions in the second buffer.

16. The system of claim 15:
wherein the buffer storage threshold comprises a first buffer storage threshold;
wherein the circuitry for detecting is further for detecting when a number of packets stored in the second buffer exceeds a second buffer storage threshold; and
5 wherein the circuitry for issuing a pause message along an output to at least a second network node is further responsive to detection by the circuitry for detecting that the number of packets stored in the second buffer exceeds the second buffer storage threshold.

17. The system of claim 16 and further comprising the second network node.

18. The system of claim 16 wherein the second network node is directly-connected to the first network node.

19. The system of claim 1:

and further comprising a plurality of network nodes;

wherein the plurality of network nodes comprise the first network node; and

wherein each network node in the plurality of network nodes comprises:

5 an input for receiving a packet;

a buffer, coupled to the input and for storing the packet;

circuitry for detecting when a number of packets stored in the buffer exceeds a buffer storage threshold; and

10 circuitry, responsive to a detection by the circuitry for detecting that the number of packets stored in the buffer exceeds the buffer storage threshold, for issuing a pause message along an output to a different network node in the plurality of network nodes;

15 wherein the pause message indicates a message ingress address and a message egress address, the message ingress address and the message egress address corresponding to a network ingress address and a network egress address in a congestion-causing packet received by the first network node; and

20 wherein the pause message commands the different network node to discontinue, for a period of time, transmitting to the network node that issued the pause message, any packets that have the message ingress address and the message egress address.

20. The system of claim 1 wherein the network node comprises a Metro Ethernet network node.

21. The system of claim 1 wherein the first network node comprises an edge node in a Metro Ethernet network and has a Metro Ethernet network address, and further comprising circuitry for coupling the packet to additional information, the additional information comprising the Metro Ethernet network address of the edge node.

22. The system of claim 21:

wherein the Metro Ethernet network address of the edge node comprises an ingress address relative to the Metro Ethernet network; and

wherein the additional information further comprises an egress Metro Ethernet
5 network address relative to the Metro Ethernet network.

23. A method of operating a network system, comprising:
at a first network node, the steps of:
- receiving a plurality of packets at an input;
 - storing the plurality of packets in a buffer, coupled to the input;
 - 5 detecting when a number of packets stored in the buffer exceeds a buffer storage threshold; and
 - responsive to detection by the detection step that the number of packets stored in the buffer exceeds the buffer storage threshold, issuing a pause message along an output to at least a second network node;
- 10 wherein the pause message indicates a message ingress address and a message egress address, the message ingress address and the message egress address corresponding to a network ingress address and a network egress address in a congestion-causing packet received by the first network node; and
- 15 wherein the pause message commands the second network node to discontinue, for a period of time, transmitting to the first network node any packets that have the message ingress address and the message egress address.

* * * * *